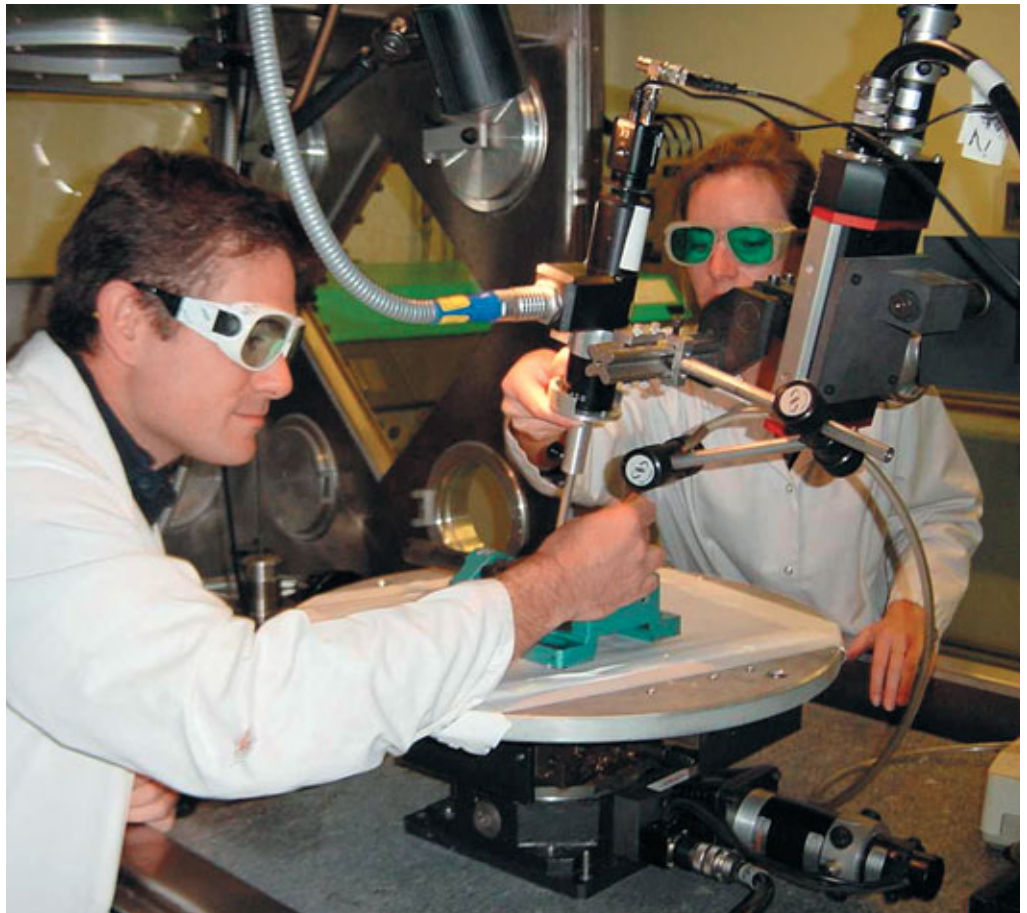


# LA MODÉLISATION COMPORTEMENTALE

*Lorsque les performances de systèmes complexes dépendent de nombreux paramètres, il est souvent difficile de déterminer quelles sont les seules "bonnes" expérimentations à réaliser. La modélisation comportementale mise en œuvre à la Direction des applications militaires du CEA constitue une aide puissante à l'optimisation de tels systèmes.*

Préparation du procédé de soudage laser YAG impulsif lors de la mise en œuvre du plan d'expériences.



CEA/DAM/CVA

## Quels objectifs ?

L'optimisation d'un système complexe pour atteindre des performances spécifiées *a priori* passe par l'identification des paramètres de fonctionnement conduisant à ces performances. Même avec l'outil de simulation le plus prédictif, s'il est possible de calculer les performances du système en fonction des paramètres de fonctionnement, l'inverse ne l'est pas. Cette recherche des meilleurs paramètres de fonctionnement et l'étude de leur sensibilité vis-à-vis de fluctuations autour de leur valeur nominale reposent sur des techniques semi-empiriques.

Pour aider à répondre à cette problématique, la Direction des applications militaires (DAM) du CEA a développé la mise en œuvre de méthodes de "modélisation comportementale" de systèmes. Une telle modélisation consiste à définir une relation mathématique entre performances et paramètres, sans connaissance *a priori* des phé-

nomènes physiques sous-jacents (encadré). Toute la difficulté est, bien entendu, d'obtenir un modèle empirique mais simple et réaliste dans le domaine d'étude. L'exploitation de ce modèle permet d'évaluer l'influence des différents paramètres, de trouver leur configuration optimale et de faire des études de sensibilité de façon immédiate et à moindre coût.

## Deux exemples d'application

### Mise au point d'un procédé de soudage laser YAG impulsif

Cette étude a été réalisée avec deux départements de la DAM implantés au centre CEA de Valduc. L'objectif était de déterminer les interactions prédominantes entre les paramètres de soudage pour garantir une forme et une pénétration du cordon de soudure définies au préalable, de modéliser de manière empirique



## La mise en œuvre de la modélisation comportementale

La modélisation comportementale de systèmes nécessite la mise en œuvre d'outils spécifiques liés au contexte de l'étude et qui puisent leur efficacité dans l'utilisation d'outils algébriques et statistiques. Deux d'entre eux peuvent être particulièrement cités : les *plans d'expériences* et les *réseaux de neurones*.

### Les plans d'expériences

Les plans d'expériences permettent de réaliser une modélisation comportementale d'un système reliant la réponse aux variables d'entrée, qu'elles soient quantitatives (modèle polynomial) ou qualitatives (modèle matriciel).

Exemple : modèle polynomial

$$Y = \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_i X_i + \alpha_{11} X_1^k + \dots + \alpha_{ij} X_i^k + \alpha_{12} X_1 X_2 + \dots + \alpha_{ij} X_i X_j + \varepsilon$$

où  $Y$  est la réponse étudiée,

$X_i$  les variables d'entrée

$\alpha_{ij}$  les coefficients à estimer

$\varepsilon$  la partie aléatoire du phénomène

$k$  le degré du polynôme pour la variable considérée.

La détermination des essais (ou calculs) à réaliser résulte de la mise en œuvre d'algorithmes d'optimisation sophistiqués<sup>(1)</sup> et répond à des critères algébriques dont le respect conditionne la qualité de l'estimation des coefficients du modèle recherché.

La minimisation du nombre d'essais et la prise en compte de leurs couplages nécessitent de faire varier conjointement plu-

sieurs facteurs d'un essai à l'autre. Quand des essais ou calculs existent déjà, il est possible d'en intégrer quelques-uns dans le plan proposé s'ils sont suffisamment pertinents (au sens mathématique décrit ci-dessus). Une exploitation statistique des résultats permet, dans le cas d'expérimentations, de dissocier les effets réels des paramètres des effets liés à la reproductibilité du processus étudié.

### Les réseaux de neurones

Ils permettent de réaliser une modélisation comportementale par des combinaisons de fonctions non linéaires. Quand les phénomènes présentent des seuils, un choix pertinent est celui de fonctions de type "sigmoïde", c'est-à-dire de forme .

Exemple : modèle Tangente hyperbolique

$$Y = \beta_1 \tanh(\alpha_{11} X_1 + \alpha_{21} X_2 + \dots + \alpha_{i1} X_i) + \beta_2 \tanh(\alpha_{12} X_1 + \alpha_{22} X_2 + \dots + \alpha_{i2} X_i) + \dots + \beta_k \tanh(\alpha_{1k} X_1 + \alpha_{2k} X_2 + \dots + \alpha_{ik} X_i)$$

où  $Y$  est la réponse étudiée

$X_i$  les variables d'entrée

$\alpha_{ij}$  et  $\beta_k$  les coefficients à estimer.

Ils présentent un grand intérêt lorsque les relations entre les réponses et les paramètres sont non linéaires (surfaces de réponse perturbées, phénomènes à seuil). La mise en œuvre de tels modèles fait appel à des algorithmes d'optimisation spécifiques<sup>(2)</sup>.

Cet outil, traditionnellement mis en œuvre pour restituer des phénomènes complexes lorsque beaucoup de données sont disponibles, peut également être utilisé à partir de peu de données, judicieusement choisies.

(1) Fedorov, recuit simulé, algorithme génétique.

(2) Gradient simple, quasi-Newton, Levenberg-Marquardt.

ces interactions et de contribuer à la mise au point des codes de calcul par **éléments finis** décrivant la thermique du soudage (voir également *Le soudage virtuel*).

Les spécialistes avaient identifié cinq paramètres supposés influents (distance de tir, fréquence et durée des impulsions, énergie et vitesse d'avance du faisceau) et présumé l'existence de nombreux couplages. Des modèles de forme polynomiale reliant cha-

cune des réponses (largeur et pénétration du cordon de soudure) à ces paramètres ont été élaborés.

La stratégie expérimentale choisie et la mise en œuvre des outils de modélisation comportementale (encadré) ont amené à réaliser 28 essais. Ceux-ci ont conduit à l'obtention de cordons de soudure assez hétéroclites compte tenu du large domaine observé (figure 1).

Ces essais ont permis de construire le modèle comportemental souhaité. Après validation par quatre essais supplémentaires, ce modèle a permis de simuler le procédé, d'étudier les effets des paramètres et de leurs couplages, de trouver le réglage optimum permettant d'obtenir la pénétration et la forme de soudure désirées et d'en démontrer la robustesse vis-à-vis de fluctuations possibles des paramètres dans leurs tolérances de fabrication.

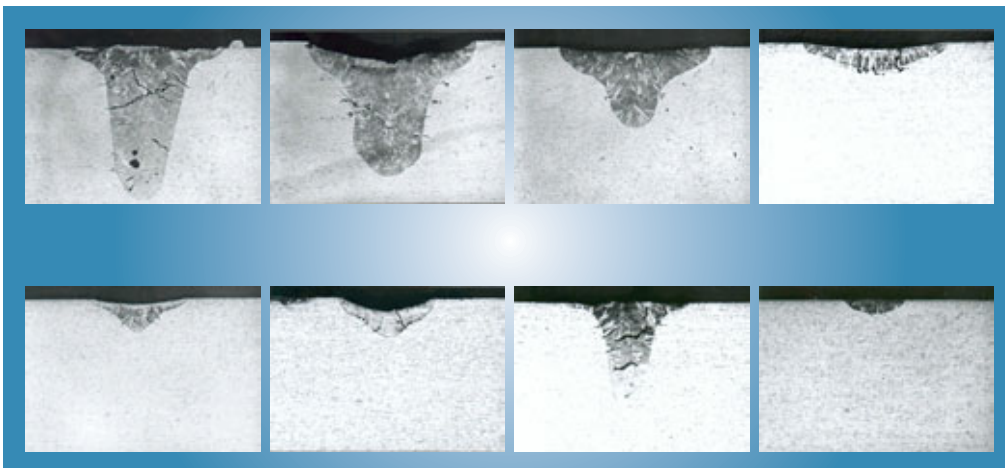
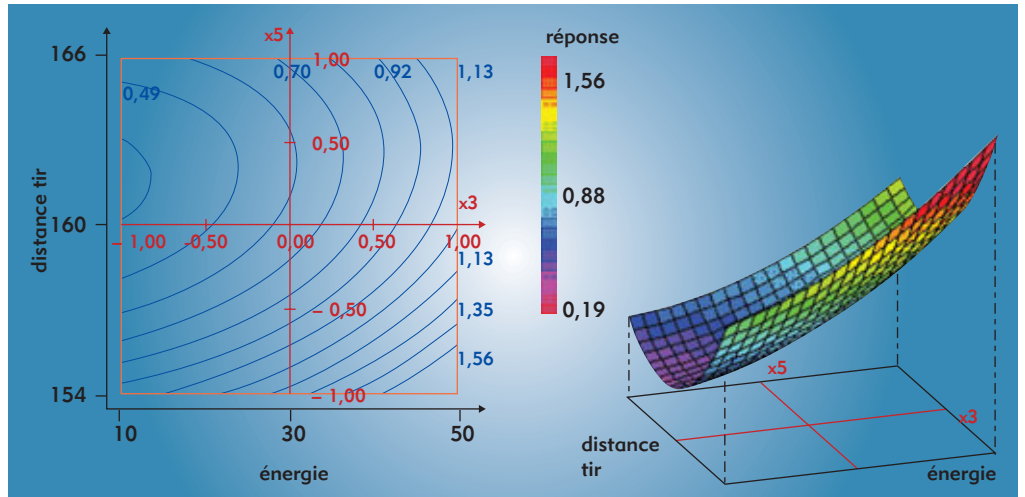


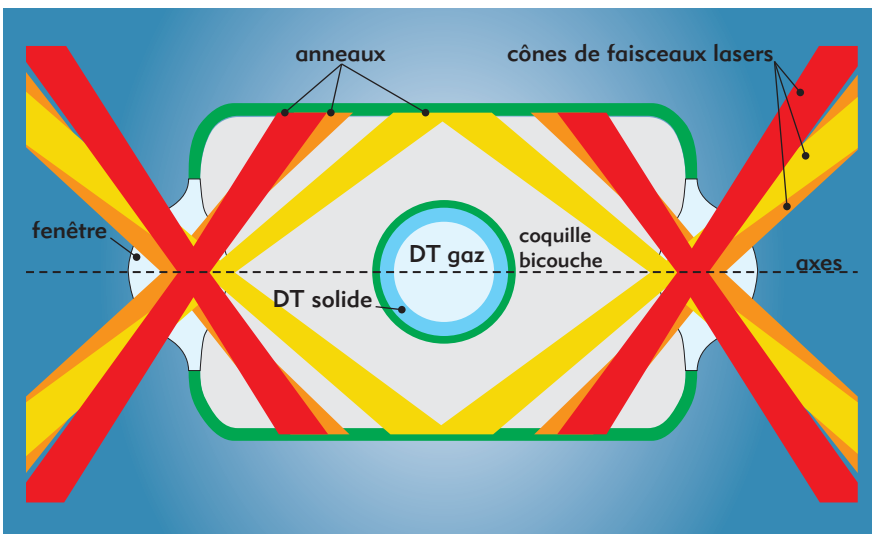
Figure 1. Exemples de coupes de cordon de soudure obtenues lors du plan d'expériences de soudage par laser YAG impulsif.

CEA/DAM

Figure 2. Exemple d'une des dix surfaces de réponse modélisant la largeur à mi-hauteur de la soudure en fonction de l'énergie du faisceau et de la distance de tir.



● ● ● ● ●  
 Figure 3. Schéma de principe d'une cible destinée au laser Mégajoule lors d'une expérience d'ignition du mélange deutérium-tritium contenu dans le microballon central.



### Évaluation de la robustesse de fonctionnement d'une cible laser lors d'une expérience d'ignition

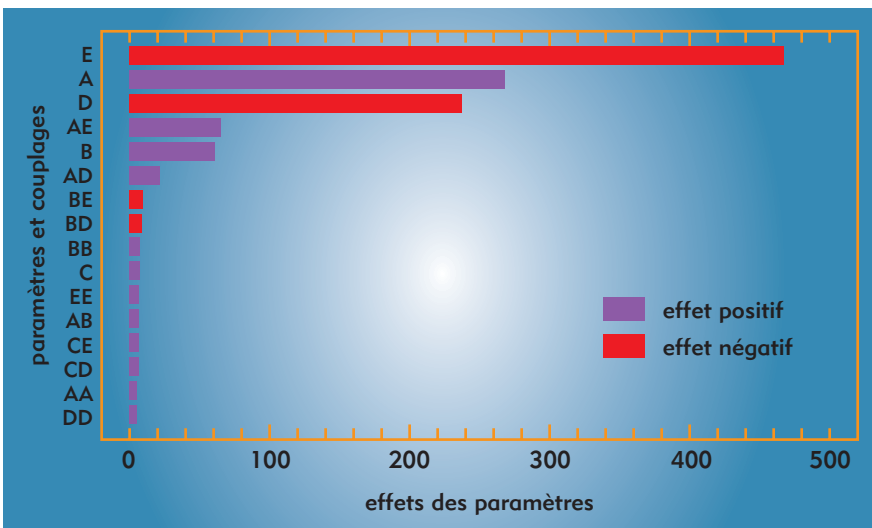
Dans cette étude, il s'agissait dans le cadre du projet Mégajoule (LMJ) (voir *Le programme Simulation : la garantie des armes sans essais nucléaires*) d'évaluer la robustesse de déformation de la cible (figure 3) aux différents défauts de pointage et d'énergie des rayons lasers incidents, lors de son implosion (2 paramètres liés à des variations du profil de la puissance appliquée, 3 autres liés aux défauts de pointage de chaque cône des faisceaux laser incidents).

Le modèle comportemental (polynomial) a été obtenu à partir de 27 calculs (au lieu de  $3^5 = 243$ ) effectués par le code utilisé pour la simulation de l'implosion. Ce modèle a permis de réaliser un très grand nombre de simulations (Monte-Carlo) afin d'évaluer cette robustesse. Malgré la complexité des phénomènes physiques en jeu (interaction laser-matière...), le modèle polynomial obtenu est représentatif de la réponse dans le domaine d'observation.

Cette modélisation a également permis de hiérarchiser les effets des paramètres (ex. : A, B, ...) et de leurs couplages (ex. : AB, AE, AD, ...) (figure 4).

● ● ● ● ●  
 Figure 4. Hiérarchisation des effets principaux et des couplages des différents paramètres intervenant sur la déformation en fonctionnement d'une cible du laser Mégajoule.

L'interprétation des phénomènes a été largement facilitée par la visualisation des surfaces de réponse en fonction des paramètres pris deux à deux (figure 2). L'accès à ces représentations aurait nécessité, sans méthode, environ 243 essais ( $3^5 : 5$  paramètres, 3 observations pour chacun).



### Des outils opérationnels

La modélisation comportementale apporte une aide précieuse à la compréhension du fonctionnement de systèmes et contribue à une conception optimisée. Actuellement, des logiciels dédiés et performants permettent sa mise en œuvre opérationnelle. Néanmoins, des recherches sont toujours en cours pour améliorer le ratio nombre d'essais (calculs) sur qualité des modèles.

Nicolas Lecler, Annie Masson  
 et Jacques Van Der Vliet  
 Direction des applications militaires  
 CEA/DAM - Ile-de-France

## Qu'est-ce qu'une simulation numérique ?

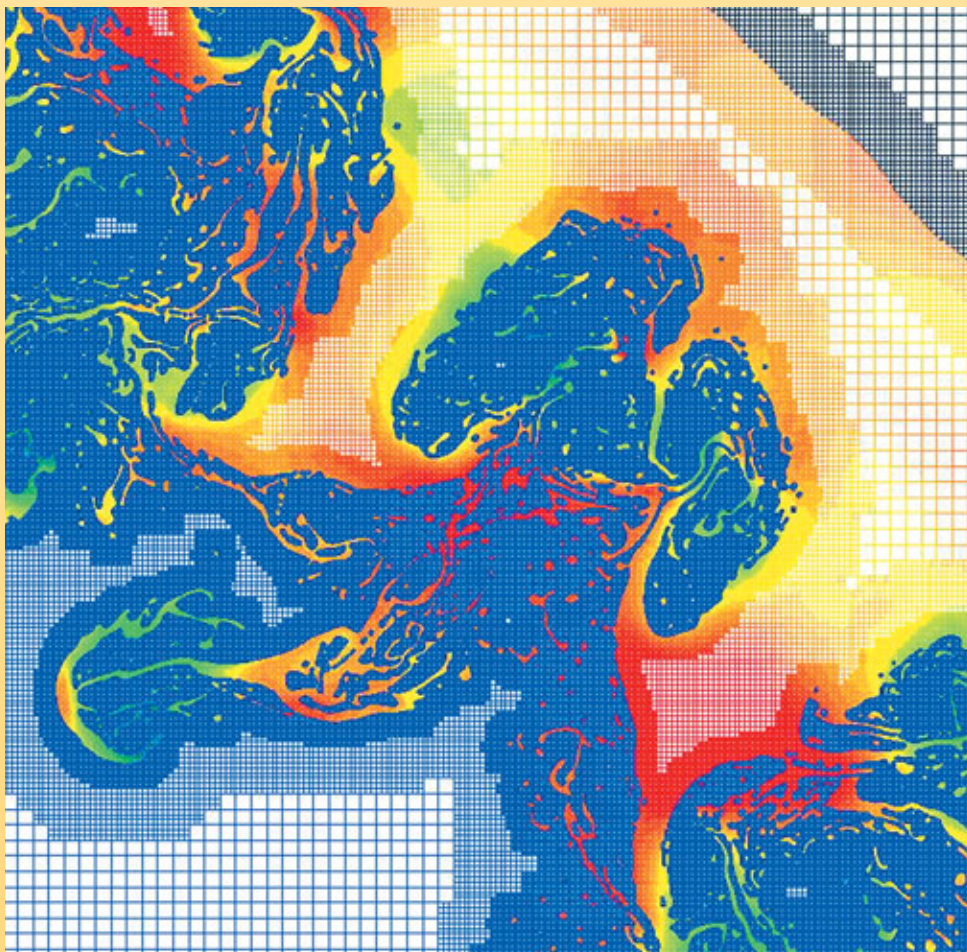
La **simulation numérique** consiste à reproduire par le calcul le fonctionnement d'un système, préalablement décrit par un ensemble de **modèles**. Elle s'appuie sur des méthodes mathématiques et informatiques spécifiques. Les principales étapes de la réalisation d'une étude par simulation numérique sont communes à de nombreux secteurs de la recherche et de l'industrie, en particulier le nucléaire, l'aérospatial ou l'automobile.

En chaque point de l'"objet" considéré, plusieurs grandeurs physiques (vitesse, température...) décrivent l'état et l'évolution du système étudié. Celles-ci ne sont pas indépendantes, mais reliées et régies par des **équations**, généralement **aux dérivées partielles**. Ces équations constituent la traduction mathématique des lois de la physique qui modélisent le comportement de l'objet. Simuler l'état de ce dernier, c'est déterminer – idéalement en tout point – les valeurs numériques de ses **paramètres**. Comme il y a un nombre infini de points, donc une infinité de valeurs à calculer, cet objectif est inaccessible (sauf dans des cas bien particuliers où l'on peut résoudre les équations de départ à l'aide de formules analytiques). Une approximation naturelle consiste donc à ne considérer qu'un nombre fini de points. Les valeurs des paramètres à calculer sont ainsi en nombre fini et les opérations nécessaires deviennent abordables grâce à l'ordinateur. Le nombre effectif de points traités dépendra bien sûr de la puissance de celui-ci : plus il sera élevé, meilleure sera finalement la description de l'objet. À la base du calcul des paramètres comme à la base de la simulation numérique, il y a donc la réduction de l'infini au fini, la **discrétisation**.

Comment opère-t-on précisément à partir des équations mathématiques du modèle ? Deux méthodes sont très souvent utilisées, respectivement représentatives des méthodes de **calcul déterministe**, qui résolvent les équations régissant les phénomènes étudiés après avoir discrétisé les variables, et des méthodes de **calcul statistique** ou **probabiliste**.

Le principe de la première, connue sous le nom de **méthode des volumes finis**, est antérieur à l'usage des ordinateurs. Chaque point de l'objet est assimilé simplement à un petit volume élémentaire (un cube par exemple), d'où le nom de *volume fini*. Un plasma, par exemple, est ainsi vu comme un ensemble ou un réseau de volumes contigus qui, par analogie avec la trame d'un tissu, sera dénommé **maillage**. Les paramètres de l'état de l'objet sont maintenant définis dans chaque maille du maillage. Pour chacune d'elles, en reformulant les équations mathématiques du modèle par des moyennes volumiques, il sera alors possible de construire des *relations algébriques* entre les paramètres de la maille et ceux de ses voisines. Au total, il y aura autant de relations que de paramètres inconnus et ce sera à l'ordinateur de résoudre le *système* de relations obtenu. Il faudra pour cela recourir aux techniques de l'**analyse numérique** et programmer des **algorithmes** spécifiques.

L'accroissement de la puissance des ordinateurs a permis d'augmenter la finesse de discrétisation, permettant de passer de quelques dizaines de mailles dans les années soixante à plusieurs dizaines de milliers dans les années quatre-vingt, à des millions dans les années quatre-vingt-dix et jusqu'à la dizaine de milliards de mailles aujourd'hui (machine Tera de la Direction



Exemple d'image d'une simulation 2D d'instabilités réalisée avec le supercalculateur Tera du CEA. Le calcul a fait appel au maillage adaptatif, qui se fait plus fin dans les zones où les phénomènes sont les plus complexes.

des applications militaires du CEA), chiffre qui devrait décupler à la fin de la décennie.

Un raffinement du maillage, le **remaillage adaptatif**, consiste à ajuster la taille des mailles en fonction des circonstances, par exemple en les rendant plus petites et plus serrées aux interfaces entre deux milieux, là où les phénomènes physiques sont les plus complexes, ou là où les variations sont les plus importantes.

La méthode des volumes finis s'applique dans des contextes physiques et mathématiques très variés. Elle autorise toute forme de maille (cube, hexaèdre, tétraèdre...) et le maillage peut être modifié durant le calcul, en fonction de critères géométriques ou physiques. Enfin, elle est aisée à mettre en œuvre dans le contexte des **ordinateurs parallèles** (encadré B, **Les moyens informatiques de la simulation numérique hautes performances**), le maillage pouvant en effet faire l'objet d'un découpage pour des calculs sur ce type de machines (exemple figure B, p. 13).

Appartiennent à la même famille la **méthode des différences finies**, cas particulier de la **méthode des volumes finis** où les côtés des mailles sont orthogonaux, et la **méthode aux éléments finis**, qui peut juxtaposer divers types de mailles. La deuxième grande méthode, dite de **Monte-Carlo**, est particulièrement adaptée pour simuler le *transport de particules*, par exemple des neutrons ou des photons d'un **plasma** (voir *Les simulations en physique des particules*). Un tel transport est en fait caractérisé par une succession d'étapes lors desquelles chaque particule peut subir différents événements (diffusion, absorption, émission...) possibles *a priori*. Les probabilités élémentaires de chacun de ces événements sont connues individuellement pour chaque particule.

Il est alors naturel d'assimiler un point du plasma à une particule. Un ensemble de particules, en nombre fini, va constituer un échantillon représentatif de l'infinité de particules du plasma, comme lors d'un sondage statistique. D'étape en étape, l'évolution de l'échantillon sera déterminée grâce à des tirages au hasard (d'où le nom de la méthode). L'efficacité de cette méthode, mise en œuvre à Los Alamos dès les années 1940, dépend bien sûr de la qualité statistique des tirages au hasard.

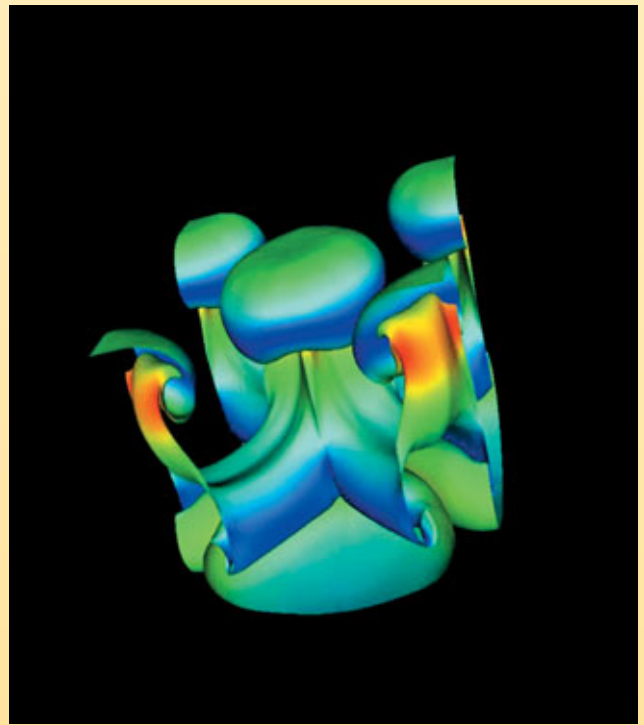
## Comment se déroule une simulation numérique ?

Il est souvent question d'*expérience numérique* pour souligner l'analogie entre la pratique d'une simulation numérique et la conduite d'une expérience de physique.

Brièvement, cette dernière utilise un dispositif expérimental, configuré selon des conditions initiales (de température, de pression...) et des paramètres de contrôle (durée de l'expérience, des mesures...). Durant l'expérience, le dispositif produit des points de mesures qui sont enregistrés. Ces enregistrements sont ensuite analysés et interprétés.

Dans une simulation numérique, le dispositif expérimental consiste en un ensemble de programmes informatiques exécutés sur des ordinateurs. Les **codes** ou **logiciels de calcul** sont la traduction, à travers des algorithmes numériques, des formulations mathématiques des modèles physiques étudiés. En amont et en aval du calcul, les *logiciels d'environnement* effectuent la gestion de plusieurs opérations complexes de préparation des calculs et de leur dépouillement.

Les données initiales de la simulation comporteront d'abord la délimitation du domaine de calcul à partir d'une représentation approchée des formes géométriques (produite par le dessin et la CAO, conception assistée par ordinateur), suivie de la discrétisation de ce



CEA  
Simulation 3D réalisée à l'aide du supercalculateur Tera installé fin 2001 au centre CEA/DAM Île-de-France à Bruyères-le-Châtel (Essonne).

Il existe pour cela des méthodes de nombres aléatoires, bien adaptées au traitement par un ordinateur.

Les méthodes des volumes finis et de Monte-Carlo ont suscité et suscitent de nombreuses études mathématiques. Ces études s'attachent notamment à préciser la convergence de ces méthodes, c'est-à-dire comment la précision de l'approximation varie avec le nombre de mailles ou de particules. Cette question est naturelle lors de la confrontation des résultats de la simulation numérique à ceux de l'expérience.

domaine de calcul sur un maillage, ainsi que les valeurs des paramètres physiques sur ce maillage et les paramètres de contrôle du bon déroulement des programmes... Toutes ces données (produites et gérées par les logiciels d'environnement) seront saisies et vérifiées par les codes. Les résultats des calculs proprement dits, c'est-à-dire les valeurs numériques des paramètres physiques, seront sauvegardés au fur et à mesure. En fait, un protocole spécifique structurera les informations produites par l'ordinateur afin de constituer une base de données numériques.

Un protocole complet organise l'échange informatique des informations requises (dimensions notamment) suivant des formats prédéfinis : *modeleur*<sup>(1)</sup>, *mailleur*<sup>(2)</sup>, *découpeur de maillage*, *code*

- (1) Le *modeleur* est un outil qui permet la création et la manipulation de points, courbes et surfaces en vue par exemple de la création d'un maillage.
- (2) Les formes géométriques d'un maillage sont décrites par des ensembles de points reliés par des courbes et des surfaces (de Bézier par exemple) qui en représentent les frontières.



de calculs, logiciel de visualisation et d'analyse. Les études de *sensibilité* des résultats (au maillage et aux modèles) font partie des "expériences" numériques.

À l'issue des calculs (résolution numérique des équations décrivant les phénomènes physiques qui se déroulent dans chaque maille), l'analyse des résultats par des spécialistes reposera sur l'exploitation de la base de données numériques. Elle comportera plusieurs étapes : extraction sélective des données (selon le paramètre physique recherché) et visualisation, extraction et transfert des données pour calculer et visualiser des diagnostics.

Le parallèle entre la conduite d'un cas de calcul, d'une expérience numérique et la conduite d'une expérience physique ne s'arrête pas là : les résultats numériques seront comparés aux résultats expérimentaux. Cette analyse comparative, effectuée sur la base de critères quantitatifs standardisés, fera appel et à l'expérience

et à l'art de l'ingénieur, du physicien, du mathématicien. Elle débouchera sur de nouvelles améliorations des modèles physiques et des programmes informatiques de simulation.

**Bruno Scheurer**

Direction des applications militaires  
CEA centre DAM-Ile de France

**Frédéric Ducros et Ulrich Bieder**

Direction de l'énergie nucléaire  
CEA centre de Grenoble

## L'exemple d'un calcul de thermohydraulique

La mise en œuvre d'un protocole de simulation numérique peut être illustrée par les travaux réalisés par l'équipe de développement du logiciel de calcul **thermohydraulique** Trio U. Ces travaux se sont déroulés dans le cadre d'une étude faite en collaboration avec l'Institut de radioprotection et de sûreté nucléaire (IRSN). L'objectif était d'obtenir des données très précises pour fournir à l'ingénieur les sollicitations en température à la paroi des composants d'un réacteur à eau sous pression dans le cas d'un accident grave impliquant une circulation naturelle turbulente de gaz chauds. Cette étude requiert la modélisation simultanée d'effets "système" à grande échelle et de phénomènes **turbulents** à petite échelle (encadré F, **Modélisation et simulation des écoulements turbulents**).

Elle débute par la définition du modèle de calcul global (figure A), suivie de la réalisation du modèle CAO et du maillage correspondant avec des logiciels du commerce (figure B). Les maillages de plus de cinq millions de mailles exigent l'utilisation de puissantes stations graphiques. Dans cet exemple, le maillage d'un générateur de vapeur (figures C et D) a été découpé pour répartir les calculs sur huit processeurs d'un ordinateur parallèle du CEA : chaque couleur symbolise une zone affectée à un processeur particulier. Les calculs, dont les conditions aux limites sont données par un calcul "système" (Icare-Cathare), produisent des résultats qu'il appartient aux spécialistes d'interpréter. En l'occurrence, les visualisations sur des stations graphiques des valeurs instantanées des champs de vitesse montrent l'impact d'un panache chaud sur la plaque tubulaire du

générateur de vapeur (coupe dans le champ de vitesses à gauche de la figure E) et la température instantanée dans la boîte à eau (à droite).

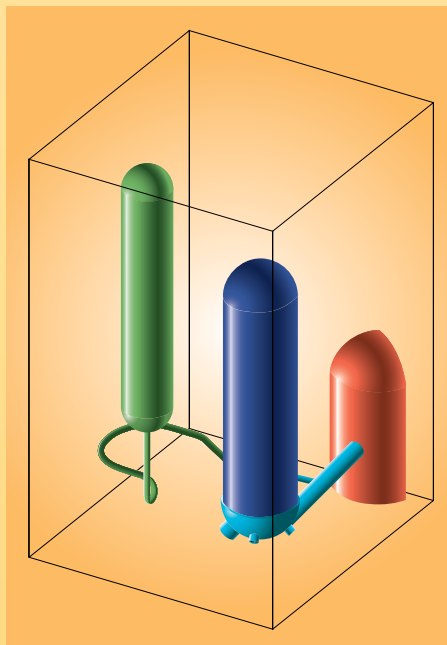


Figure A. Domaine de calcul global incluant une partie de la cuve réacteur (rouge), la conduite de sortie (branche chaude en bleu clair), le générateur de vapeur (bleu foncé) et le pressuriseur (vert).

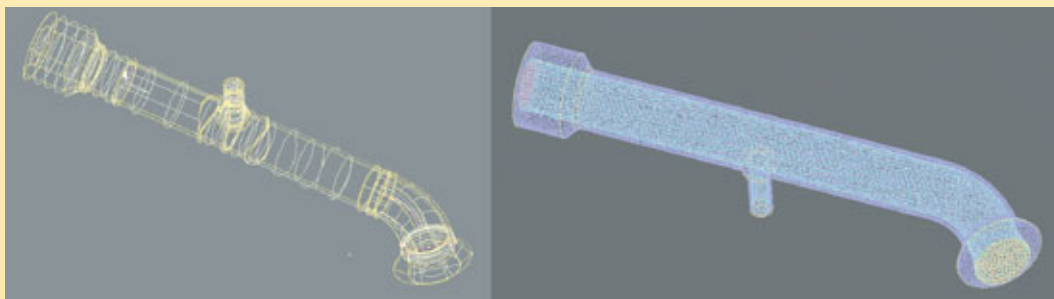
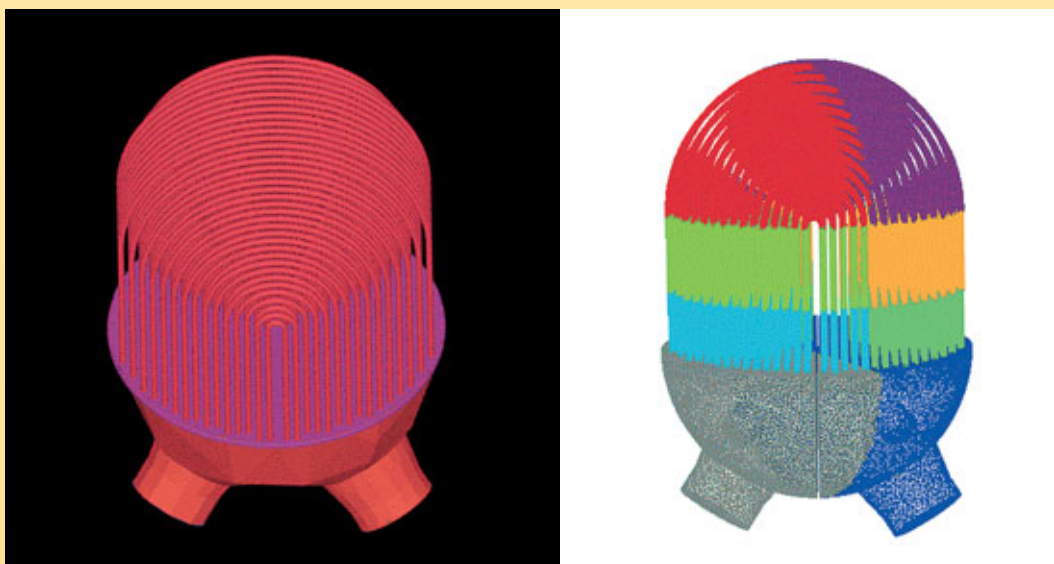


Figure B. Modèle CAO de la branche chaude en sortie de la cuve réacteur (à gauche) et son maillage non structuré (à droite).



Figures C et D.

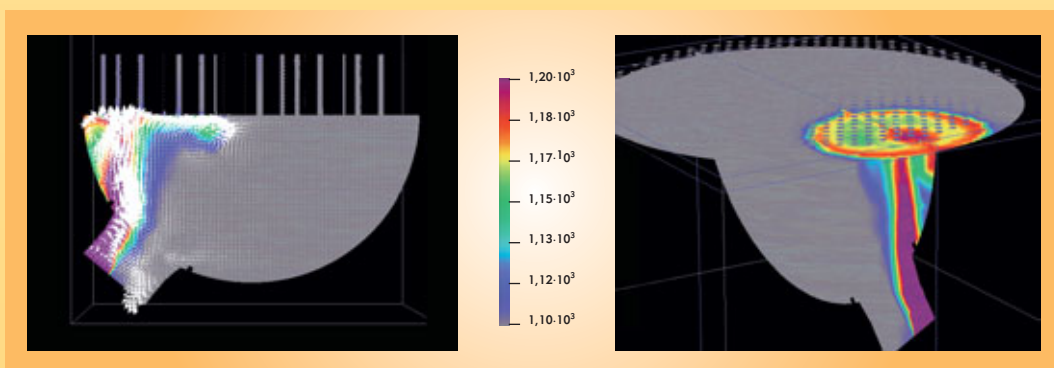


Figure E.

# Les moyens informatiques de la simulation

Effectuer des **simulations numériques** plus précises impose de mettre en œuvre des **modèles** physiques et numériques eux-mêmes plus précis portant sur des descriptions plus fines des objets simulés (encadré A, *Qu'est-ce qu'une simulation numérique ?*). Tout ceci nécessite des progrès dans le domaine des logiciels de simulation mais aussi une augmentation importante de la capacité des équipements informatiques sur lesquels ces logiciels sont utilisés.

## Processeurs scalaires et vectoriels

Au cœur de l'ordinateur, le processeur est l'unité de base qui, déroulant un programme, effectue les calculs. Il en existe deux grands types, les **processeurs scalaires** et les **processeurs vectoriels**. Les premiers exécutent des opérations portant sur des nombres élémentaires (scalaires), par exemple l'addition de deux nombres. Les seconds exécutent des opérations portant sur des ensembles de nombres (vecteurs), par exemple additionner deux à deux les nombres composant deux ensembles de 500 éléments. À ce titre, ils sont particulièrement adaptés à la simulation numérique : lors de l'exécution d'une opération de ce type, un processeur vectoriel peut fonctionner à une vitesse proche de sa performance maximale (crête). La même opération avec un processeur scalaire exige de nombreuses opérations indépendantes (opérations par composante des vecteurs) qui s'exécutent à une vitesse bien inférieure à sa vitesse crête. L'avantage principal des processeurs scalaires est leur prix : il s'agit de microprocesseurs généralistes dont les coûts de conception et de fabrication peuvent être amortis sur de larges marchés.

## Forces et contraintes du parallélisme

Les processeurs récents permettent de hautes performances, d'une part en utilisant une fréquence de fonctionnement plus élevée, d'autre part en cherchant à exécuter en même temps

plusieurs opérations : c'est un premier niveau de **parallélisme**. L'accélération de la fréquence est limitée par l'évolution de la technologie micro-électronique, tandis que les dépendances entre instructions à exécuter par le processeur limitent le parallélisme possible. La mise en œuvre simultanée de plusieurs processeurs constitue un second niveau de parallélisme, qui permet d'obtenir des performances accrues à condition de disposer de programmes capables d'en tirer parti. Alors que le parallélisme au niveau des processeurs est automatique, celui *entre processeurs* dans un ordinateur parallèle est à la charge du programmeur, qui doit découper son programme en morceaux indépendants et prévoir entre eux les communications nécessaires. On procède souvent par un découpage du domaine sur lequel porte le calcul, chaque processeur étant chargé de simuler le comportement d'un domaine, et par l'établissement de communications régulières entre processeurs afin de garantir la cohérence d'ensemble du calcul. Pour obtenir un programme parallèle efficace, il faut s'assurer de l'équilibrage de charge entre processeurs et chercher à limiter le coût des communications.

## Les différentes architectures

Les équipements informatiques ont différentes fonctions. À partir de son ordinateur de travail sur lequel il prépare ses calculs et en analyse les résultats, l'utilisateur accède à des moyens de calcul, de stockage, et de visualisation partagés, mais beaucoup plus puissants que les siens propres. L'ensemble de ces équipements sont reliés par des réseaux informatiques permettant de faire circuler les informations entre eux avec des débits compatibles avec le volume de données produites, pouvant atteindre 1 **téraoctet** ( $1 \text{ To} = 10^{12}$  octets) de données pour une seule simulation.

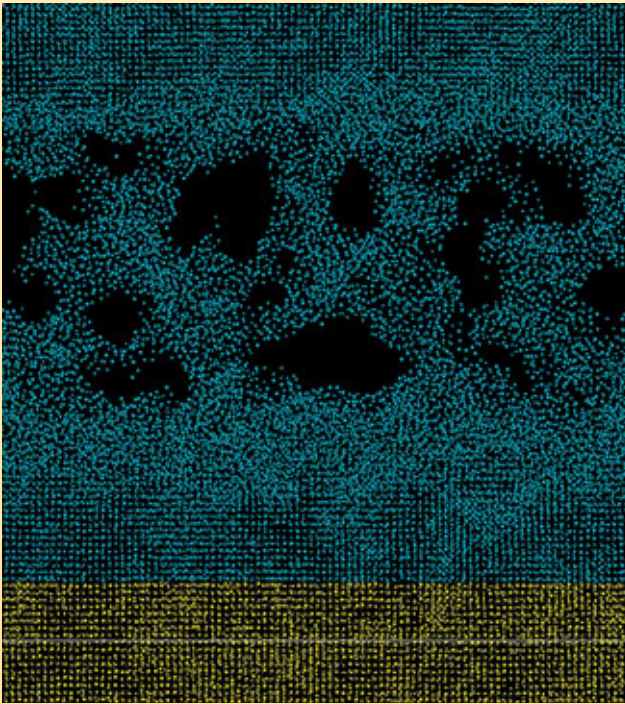
Les grands équipements de calcul sont généralement appelés **superordinateurs**. Ils atteignent aujourd'hui des puissances qui se chiffrent en **téraflops** (1 Tflops =  $10^{12}$  opérations de calcul par seconde).

Il existe aujourd'hui trois grands types de superordinateurs : les superordinateurs vectoriels, les grappes de mini-ordinateurs à mémoire partagée et les grappes de PC (l'ordinateur que chacun possède chez soi). Le choix entre ces architectures dépend largement des applications et de l'utilisation visées. Les superordinateurs vectoriels disposent de processeurs très performants mais dont il est difficile d'augmenter la puissance en ajoutant des processeurs. Les grappes de PC sont peu coûteuses mais mal adaptées à des environnements où de nombreux utilisateurs font beaucoup de calculs très gourmands en puissance machine, en mémoire et en entrées-sorties.

Ce sont ces considérations qui ont en particulier conduit la Direction des applications militaires (DAM) du CEA à choisir pour son programme simulation (voir *Le programme Simulation : la garantie des armes sans essais nucléaires*) les architectures de type grappe de mini-ordinateurs à mémoire partagée, encore appelées **clusters de SMP** (Symmetric Multi-Processor). Un tel système utilise comme brique de base un mini-ordinateur com-



Installée en décembre 2001 au CEA (centre DAM-Ile de France) et conçue par Compaq (devenue depuis HP), la machine Tera a pour élément de base un mini-ordinateur à 4 processeurs Alpha à 1 GHz partageant une mémoire de 4 Go et fournissant une puissance totale de 8 Gflops. Ces éléments de base sont interconnectés par un réseau rapide conçu par la société Quadrics. Une opération de synchronisation sur l'ensemble des 2 560 processeurs s'effectue en moins de 25 microsecondes. Le système de fichiers global offre un espace de stockage de 50 téraoctets pour les entrées-sorties avec une bande passante agrégée de 7,5 Go/s.



CEA

Les calculateurs parallèles sont adaptés aux méthodes numériques basées sur des maillages (encadré A, **Qu'est-ce qu'une simulation numérique ?**) mais aussi au traitement de calculs ab initio comme cette simulation par dynamique moléculaire de l'endommagement par choc de deux plaques de cuivre à 1 km/s (voir La simulation des matériaux). Le système considéré est constitué de 100 000 atomes de cuivre représentant un parallélépipède de section carrée (0,02 µm de côté) à densité normale. Les atomes interagissent suivant un potentiel EAM (embedded atom potential) pendant 4,6 picosecondes. Le calcul, effectué sur 18 processeurs du supercalculateur Tera de Bruyères-le-Châtel à l'aide du logiciel Stamp développé au CEA, a représenté une dizaine de minutes de temps "utilisateur" (calcul réalisé par B. Magne). Des tests impliquant jusqu'à 64 millions d'atomes ont été réalisés, mobilisant 256 processeurs pendant une centaine d'heures.

portant plusieurs microprocesseurs qui partagent une mémoire commune (figure). Ces mini-ordinateurs étant largement diffusés dans des domaines variés allant de la banque au serveur web

en passant par les bureaux d'études, ils offrent un excellent rapport performance/prix. Ces "briques" de base (encore appelées *nœuds*) sont reliées entre elles par un réseau d'interconnexion hautes performances : la puissance cumulée de plusieurs centaines de ces "briques" peut atteindre plusieurs téraflows. On parle alors d'**ordinateur massivement parallèle**.

Cette puissance peut être disponible pour une seule application parallèle utilisant toutes les ressources du supercalculateur mais aussi pour de multiples applications indépendantes, parallèles ou non, utilisant chacune une partie des ressources.

Si la caractéristique mise en avant pour décrire un supercalculateur est en général sa puissance de calcul, il ne faut pas négliger l'aspect entrées-sorties. Ces machines capables d'effectuer des simulations de grande taille doivent disposer de systèmes de disques avec des capacités et des performances adaptées. Dans les *clusters* de SMP, chaque mini-ordinateur dispose d'un espace disque local. Il n'est néanmoins pas judicieux d'utiliser celui-ci pour les fichiers utilisateurs, ce qui obligerait l'utilisateur à explicitement déplacer ses données entre les différentes phases de ses calculs. Pour cette raison, il est important de disposer d'un espace disque accessible par l'ensemble des mini-ordinateurs du supercalculateur. Cet espace est en général constitué de batteries de disques reliées à des nœuds dont la fonction principale est de les gérer. Comme pour le calcul, c'est le parallélisme pour les entrées-sorties qui permet d'offrir des performances élevées. Il faut, pour ce faire, disposer de systèmes de fichiers globaux parallèles permettant un accès rapide et sans contraintes à l'espace disque partagé. Offrant des puissances de calcul considérables, les *clusters* de SMP posent néanmoins plusieurs défis. Parmi les plus importants, outre la programmation de logiciels de simulation capables de tirer parti du grand nombre de processeurs, il faut mettre au point des systèmes d'exploitation et les logiciels associés compatibles avec de telles configurations et tolérants vis-à-vis des pannes.

**François Robin**

Direction des applications militaires  
CEA centre DAM-Ile de France

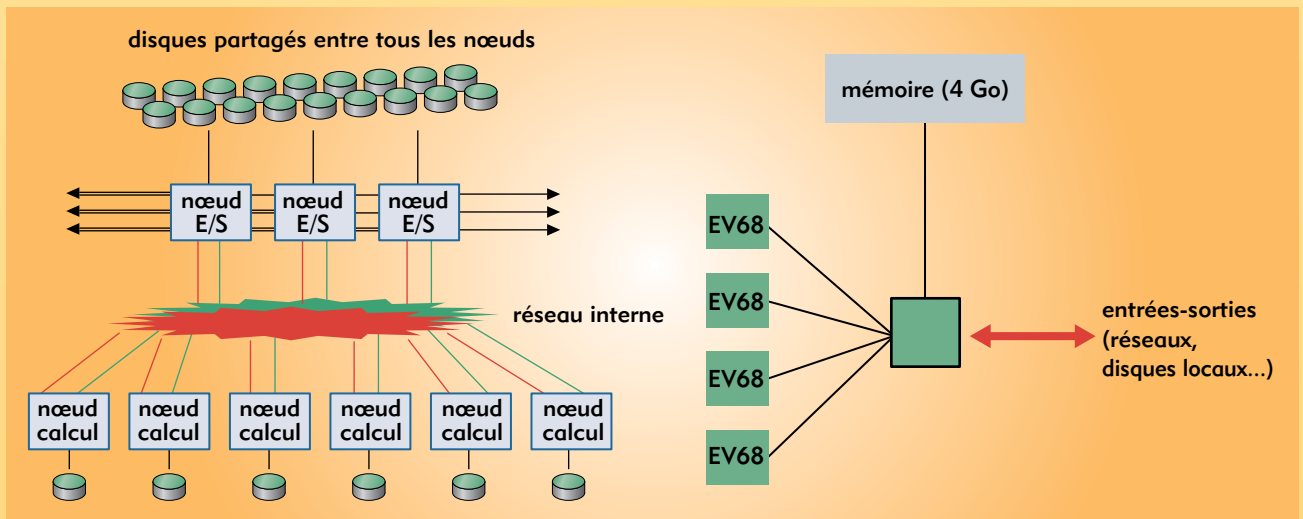


Figure. Architecture d'une machine du type "cluster de SMP". À gauche, l'architecture générale (E/S = entrée/sortie), à droite celle d'un nœud avec quatre processeurs Alpha EV68 cadencés à 1 GHz.

# Modélisation et simulation des écoulements turbulents

La **turbulence**, ou l'agitation de l'écoulement dit turbulent, se développe dans la plupart des écoulements qui conditionnent notre environnement immédiat (rivières, océan, atmosphère). Elle se révèle être aussi un, sinon le, paramètre dimensionnant dans un bon nombre d'écoulements industriels (liés à la production ou la conversion d'énergie, à l'aérodynamique...). Il n'est donc pas étonnant que soient entrepris des efforts visant sa prédiction – fût-elle encore imprécise – surtout lorsqu'elle se trouve combinée à des phénomènes qui la compliquent : stratification, combustion, présence de plusieurs phases... C'est que, paradoxalement, même s'il est possible d'anticiper la nature turbulente d'un écoulement et même, d'un point de vue théorique, de dégager certaines caractéristiques communes et apparemment universelles aux écoulements turbulents<sup>(1)</sup>, leur prédiction dans

des cas précis reste délicate. Celle-ci doit en effet prendre en compte l'importante gamme d'échelles spatiales et temporelles<sup>(2)</sup> impliquées dans tout écoulement de ce type.

Les chercheurs ne sont pourtant pas démunis, aujourd'hui, pour aborder ce problème. En premier lieu, les équations qui régissent l'évolution spatio-temporelle des écoulements turbulents (équations de Navier-Stokes<sup>(3)</sup>) sont connues. Leur résolution complète, dans des cas très favorables, a conduit à des descriptions prédictives. Mais l'emploi systématique de cette méthode de résolution se heurte à deux difficultés rédhibitoires : d'une part, il nécessiterait la connaissance complète et simultanée de toutes les variables attachées à l'écoulement et des forçages s'exerçant sur lui<sup>(4)</sup> et, d'autre part, il mobiliserait des moyens de calculs irréalistes pour encore des décennies.

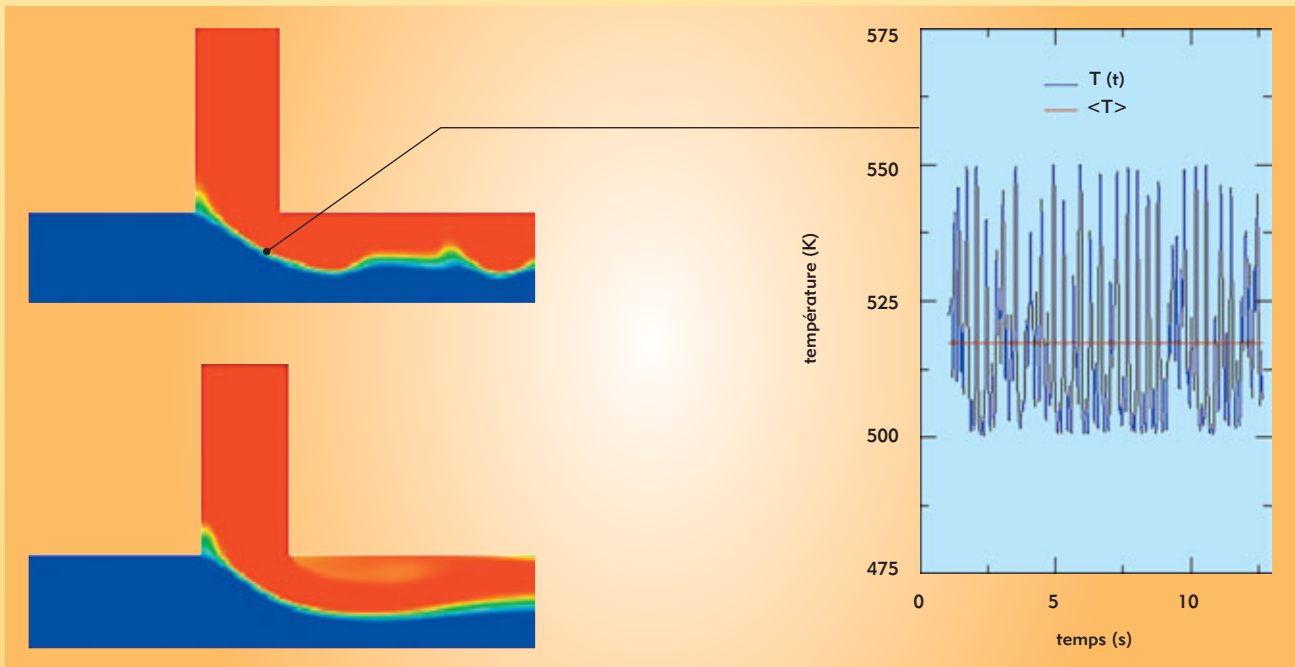


Figure. Champ de température instantané (haut) et moyenné (bas) dans une situation de mélange. La courbe donne l'historique de la température en un point : valeur instantanée fluctuante en bleu et moyenne en rouge (d'après la thèse d'Alexandre Chatelain [DEN/DTP/SMTH/LDTA]).

Il faut donc se résoudre, en s'appuyant sur le caractère fluctuant dû à l'agitation turbulente, à définir et utiliser des moyennes. Une des approches les plus répandues consiste à aborder le problème sous un angle statistique. Les moyennes d'ensemble de vitesse, de pression, de température... dont la distribution caractérise l'écoulement turbulent sont définies comme les variables principales de l'écoulement qu'on cherche à qualifier par rapport à ces moyennes. Ceci conduit à une décomposition du mouvement (dite de Reynolds) en champs moyen et fluctuant, ce dernier mesurant l'écart instantané et local entre chaque grandeur réelle et sa moyenne (figure). Ces fluctuations représentent la turbulence et couvrent une partie importante du spectre de Kolmogorov<sup>(1)</sup>.

Cette opération réduit considérablement le nombre de degrés de liberté du problème et le rend « manipulable » informatiquement. Elle comporte aussi de nombreuses difficultés : il faut tout d'abord constater que, précisément en raison des non-linéarités des équations du mouvement, toute moyenne fait surgir des termes nouveaux et inconnus qu'il faut estimer. En fermant la porte à la description complète et déterministe du phénomène, on ouvre celle de la modélisation, c'est-à-dire à la représentation des effets de la turbulence sur les variables moyennes.

Beaucoup de progrès ont été accomplis depuis les premiers modèles (Prandtl, 1925). Les modélisations n'ont cessé d'évoluer vers plus de complexité, se basant sur le fait généralement vérifié que toute nouvelle extension permet de conserver les propriétés antérieurement acquises. Il faut aussi constater que, même

si de nombreux développements remettent en avant la nécessité de traiter les écoulements en respectant leur caractère *instationnaire*, les modélisations les plus populaires ont été développées dans le cadre des écoulements *stationnaires*, pour lesquels on n'accède donc qu'à une représentation de la moyenne temporelle de l'écoulement : dans le modèle mathématique final, les effets de la turbulence proviennent ainsi intégralement de la modélisation.

Il est également remarquable que, malgré de nombreux travaux, aucune modélisation n'est aujourd'hui capable de rendre compte de l'intégralité des phénomènes qui influencent la turbulence ou sont influencés par elle (transition, instationnarité, stratification, compression, etc.). Ce qui semble pour l'instant empêcher les modélisations statistiques de nourrir une ambition d'universalité.

Malgré ces limitations, la plupart des modélisations statistiques courantes sont maintenant disponibles dans les codes commerciaux et les outils des industriels. Il n'est pas possible de prétendre qu'elles permettent des calculs prédictifs dans toute situation. Leur précision est variable, offrant des résultats utiles pour l'ingénieur dans des situations maîtrisées et favorables (prédiction de la trainée avec une précision de 5 % à 10 % d'erreur [parfois mieux] sur certains profils), mais parfois faux dans des situations qui se révèlent, après coup, en dehors du champ de validité du modèle. Tout emploi maîtrisé d'une modélisation repose donc sur une qualification particulière au type d'écoulement à traiter. Des modélisations alternatives, répondant au besoin d'une plus grande précision sur des gammes d'échelles spatiales et temporelles plus étendues et donc basées sur un opérateur de "moyenne" d'une nature différente, sont actuellement en développement et représentent des voies nouvelles.

Le paysage des modélisations de la turbulence est aujourd'hui très complexe et l'unification des points de vue et des divers concepts de modélisation est une gageure. La tentation de l'universalité des modélisations reste donc hors de propos. Leur mise en œuvre réelle relève la plupart du temps de compromis généralement guidés par le savoir-faire de l'ingénieur.

(1) On peut faire référence à la répartition spectrale de l'énergie cinétique turbulente, connue comme le "spectre de Kolmogorov", qui illustre de manière très simple la hiérarchie des échelles, des grandes échelles porteuses d'énergie aux échelles de plus en plus petites et de moins en moins énergétiques.

(2) Cette étendue est le résultat des non-linéarités des équations du mouvement qui donne naissance à une gamme étendue d'échelles spatiales et temporelles. Cette gamme est une fonction croissante du nombre de Reynolds,  $Re$ , mesurant le rapport entre force d'inertie et force visqueuse.

(3) L'hypothèse selon laquelle la résolution complète des équations de Navier-Stokes permet la simulation de la turbulence est généralement admise, tout du moins dans la gamme des écoulements sans choc.

(4) Il s'agit d'un problème régi par des conditions initiales et aux limites.